

基于四维帧内预测和4D-MDCT的彩色视频编码

赵岩, 陈贺新

(吉林大学通信工程学院通信新技术重点实验室, 长春 130012)

摘要: 将彩色视频表示在四维矩阵的数学模型中, 并利用四维矩阵离散余弦变换(4D-MDCT)统一处理彩色视频的多个帧是视频编码算法的一个新思路。该文基于4D-MDCT并结合四维帧内预测技术提出了一种彩色视频编码的新方法。实验结果表明, 该方法能够进一步提高基于4D-MDCT的彩色视频压缩编码的性能。

关键词: 四维帧内预测; 四维矩阵; 离散余弦变换; 视频编码

Color Video Coding Based on 4D Intra Prediction and 4D-MDCT

ZHAO Yan, CHEN Hexin

(Key Lab of New Communication Technique, College of Communication Engineering, Jilin University, Changchun 130012)

【Abstract】 It is a novel idea of video coding algorithm that represents color video in the mathematic model of four-dimensional matrix and encodes multiple frames of color video using four-dimensional matrix discrete cosine transform(4D-MDCT). A new color video coding method which combines 4D-MDCT and four dimensional intra prediction is proposed. The experimental results show that the proposed algorithm can improve the performance of color video compression based on 4D-MDCT further.

【Key words】 Four-dimensional intra prediction; Four-dimensional matrix; Discrete cosine transform; Video coding

在 MPEG-2 及其它传统国际视频压缩编码标准, 如 H.261、H.263、MPEG(运动图像专家组)-1 以及 20 世纪 90 年代发展起来的一些基于分形理论, 小波分析理论等的编码方法中, 一般都是先将彩色图像或视频的 R、G、B 数据转换成 Y、U、V 数据, 然后将灰度图像的压缩编码方法或其改进方法独立地应用于各个分量上, 没能充分利用 R、G、B 各颜色分量间存在的潜在相关性。据此, 文献[1]首次提出了矩阵变换理论并将其应用于彩色图像压缩。文献[2]进一步发展了三维矩阵理论, 并运用三维矩阵变换方法将彩色图像的 R、G、B 3 帧统一在一个模型中进行编码, 获得了优于 JPEG(联合图像专家组)标准方法的信噪比和压缩比性能。文献[3]将彩色视频用四维矩阵表示, 并利用四维矩阵离散余弦变换(4D-MDCT), 把对彩色视频多个帧的处理统一起来, 转化为对空间四维变换的数学处理, 全面去除彩色视频中各灰度图像的像素之间, 彩色空间的 R、G、B 3 帧之间, 以及视频的连续帧之间的相关性, 从而获得了优于 MPEG-2 中采用的基于二维 DCT 和运动补偿技术的编码性能。

本文将 4D-MDCT 与四维帧内预测技术相结合。首先对彩色视频利用帧内预测技术去除空间相关性, 进而对预测误差信号做 4D-MDCT 和矢量量化, 从而进一步改善了文献[3]中编码方法的压缩性能。

1 四维帧内预测

帧内预测是去除视频图像内相邻宏块间像素空间相关性的技术。本文中采用 4 种 $16 \times 16 \times 3 \times 3$ 的四维帧内预测模式, 这 4 种模式分别为垂直预测模式、水平预测模式、直流预测模式和平面预测模式, 各模式的行列剖面图如图 1(a)~图 1(d)所示。图 1 中 H 为待预测宏块的上面相邻宏块中最后一行的像素值, V 为左边相邻宏块内最右边一列的像素值。垂直预测模式中, 待预测宏块由 H 外推形成; 水平预测模式中, 待

预测宏块由 V 外推形成; 直流预测模式中, 待预测宏块的所有像素值均为 H 和 V 的均值; 平面预测模式中, 待预测宏块中的像素值由 H、V 和 X 的线性“平面”函数生成。

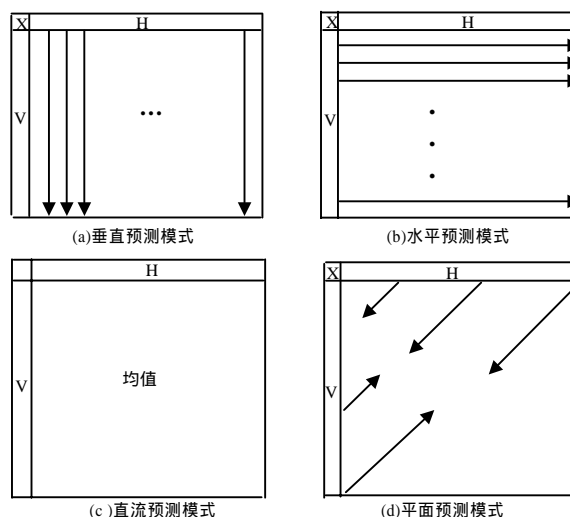


图 1 4 种帧内预测模式的行列剖面

设 $P(i, j, k, l)$ 表示待预测宏块的预测像素值, $F(i, -1, k, l)$ 表示待预测宏块的上面相邻宏块中最后一行的像素值(如图 1 中的 H), $F(-1, j, k, l)$ 表示待预测宏块的左边相邻宏块内最右边一列的像素值(如图 1 中的 V), $F(-1, -1, k, l)$ 表示待预测宏

基金项目: 国家自然科学基金资助项目(60372060); 国家自然科学基金资助重大国际合作项目(60211130507); 吉林大学青年基金资助项目(2003QN038)

作者简介: 赵岩(1971-), 女, 副教授、博士, 主研方向: 多维矩阵理论, 视频信号处理, 图像通信等; 陈贺新, 教授、博士、博导
收稿日期: 2005-11-25 **E-mail:** zhao@email.jlu.edu.cn

块的左上角相邻宏块的最后一个像素值(如图 1 中的 X), 其中, $i = 0, 1, \dots, 15; j = 0, 1, \dots, 15; k = 0, 1, 2; l = 0, 1, 2$ 。

1.1 垂直预测模式

只有当存在上面相邻宏块, 即相邻像素 $F(i-1, k, l)$ 存在时, 才存在垂直预测模式, 此时 $P(i, j, k, l) = F(i-1, k, l)$ 。

1.2 水平预测模式

只有当存在左边相邻宏块, 即相邻像素 $F(-1, j, k, l)$ 存在时, 才存在水平预测模式, 此时 $P(i, j, k, l) = F(-1, j, k, l)$ 。

1.3 直流预测模式

不管待预测宏块是否有相邻宏块都存在直流预测模式。当相邻像素 $F(-1, j, k, l)$ 和 $F(i-1, k, l)$ 都存在时, 有

$$P(i, j, k, l) = \left[\sum_{x=0}^{15} F(x, -1, k, l) + \sum_{y=0}^{15} F(-1, y, k, l) + 16 \right] \gg 5$$

当相邻像素 $F(-1, j, k, l)$ 存在而 $F(i-1, k, l)$ 不存在时, 有

$$P(i, j, k, l) = \left[\sum_{y=0}^{15} F(-1, y, k, l) + 8 \right] \gg 4$$

当相邻像素 $F(i-1, k, l)$ 存在而 $F(-1, j, k, l)$ 不存在时, 有

$$P(i, j, k, l) = \left[\sum_{x=0}^{15} F(x, -1, k, l) + 8 \right] \gg 4$$

如果相邻像素 $F(-1, j, k, l)$ 和 $F(i-1, k, l)$ 都不存在, 则

$$P(i, j, k, l) = 128$$

1.4 平面预测模式

只有当所有相邻像素 $F(-1, j, k, l)$ 、 $F(i-1, k, l)$ 和 $F(-1, -1, k, l)$ 都存在时, 才有平面预测模式。此时

$$P(i, j, k, l) = [a + b(i-7) + c(j-7) + 16] \gg 5, \text{ 其中}$$

$$a = 16 * [P(-1, 15, k, l) + P(15, -1, k, l)]$$

$$b = (5 * H + 32) \gg 6$$

$$c = (5 * V + 32) \gg 6$$

$$H = \sum_{x=1}^8 x [P(7+x, -1, k, l) - P(7-x, -1, k, l)]$$

$$V = \sum_{y=1}^8 y [P(-1, 7+y, k, l) - P(-1, 7-y, k, l)]$$

又设 $R(i, j, k, l)$ 为待预测宏块的实际像素值, SAE 表示某种预测模式下的预测误差, 则有

$$SAE = \sum_{i=0}^{15} \sum_{j=0}^{15} \sum_{k=0}^2 \sum_{l=0}^2 |R(i, j, k, l) - P(i, j, k, l)|$$

即 SAE 为四维帧内预测误差宏块中各误差的绝对值和。使 SAE 最小的预测模式则为待预测宏块的预测模式, 预测值为相应模式下的 $P(i, j, k, l)$ 。

2 基于四维帧内预测和 4D-MDCT 的视频编码

基于四维帧内预测和 4D-MDCT 的视频编码方案如图 2。

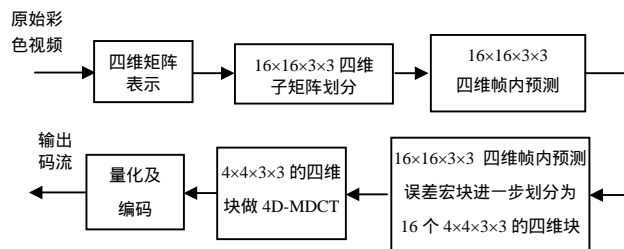


图 2 基于四维帧内预测和 4D-MDCT 的视频编码方案

3 实验结果与分析

本文对图像大小为 176×144 (QCIF 格式) 的原始彩色视频序列“News”的前 12 帧利用提出的方法进行了压缩编码,

并将编码结果同文献[3]的方法进行了对比分析。为方便对比, 本文的方法也采用了 $4 \times 4 \times 3 \times 3$ 的子矩阵变换。表 1 列出了平均 PSNR 的对比结果。图 3 给出了压缩比特率约为 139kbps 时各帧的亮度(Y)PSNR 对比曲线。为了比较主观图像质量, 图 4 (a)、图 4 (b)和图 4 (c)分别给出了第 5 帧原图像和采用两种编码方法的恢复图像。

表 1 平均 PSNR 的对比结果

本文的方法		文献[3]的方法	
比特率(kbps)	Y-PSNR (dB)	比特率(kbps)	Y-PSNR (dB)
138.98	33.64	139.22	32.57
120.86	29.79	123.75	28.68

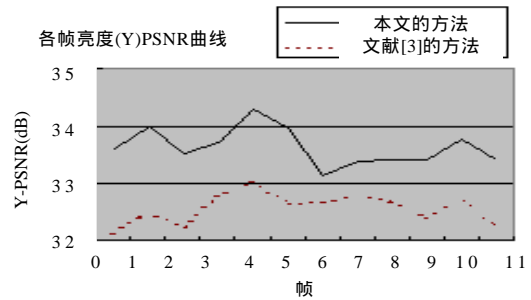


图 3 各帧的亮度(Y)PSNR 对比曲线



(a)第 5 帧原图像



(b)采用本文方法的第 5 帧恢复图像



(c)采用文献[3]方法的第 5 帧恢复图像

图 4 第 5 帧原图像和恢复图像

4 结论

本文将四维帧内预测技术引入基于 4D-MDCT 的彩色视频编码。四维帧内预测去除了视频图像内相邻宏块间像素的空间相关性, 在此基础上进行基于 4D-MDCT 的彩色视频编码, 从而实现更高的压缩性能。实验结果表明, 无论从客观评价标准还是从主观恢复效果上, 本文方法均优于文献[3]的方法, 进一步提高了彩色视频的压缩性能。

参考文献

- 1 朱艳秋, 陈贺新, 戴逸松. 彩色图像压缩矩阵变换编码[J]. 电子学报, 1997, 25(7): 16-21.
- 2 桑爱军, 陈贺新. 三维矩阵彩色图像 WDCT 压缩编码[J]. 电子学报, 2002, 30(4): 594-597.
- 3 赵 岩, 陈贺新. 彩色视频的四维矩阵离散余弦变换编码[J]. 中国图像图形学报, 2003, 18(6): 620-624.